

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OK protein - nucleic search, using frame_plus.p2n model

Run on: December 9, 2002, 12:23:04 ; Search time 2719 Seconds
(without alignments)
3574.968 Million cell updates/sec

Title: US-09-765-034-2
Perfect score: 1747
Sequence: 1 MCGIMANNATCKNMLAAEA.....KSLTFSRRAHLLSFRK 334

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Zgapop 6.0 , Zgapext 7.0
Delep 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgml/1/USPTO.spool/US09765034/runat_05122002.132015.21439/app_query.fasta.1.519
-DB=GenEmbl -GENE=fastap -SUPER=p2n.rge -MINMATCH=0.1 -LOOPCT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45
-DOCLIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTPMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09765034.CCGN.1.1.1616.grunat.05122002.132015.21439 -KCPQ=6 -ICPU=3
-NO.XLPHY -NO.MAP -LARGOQUERY -NEG.SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -YGAPOP=6 -YGAPEXT=7
-ZGAPOP=10 -ZGAPEXT=0.5 -DELEP=6 -DELEXT=7

Database : GenEmbl:*
1: gb_ba:*
2: gb_hlg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*

29: em_vl:*
30: em_hlg_hum:*
31: em_hlg_inv:*
32: em_hlg_other:*
33: em_hlg_mus:*
34: em_hlg_pln:*
35: em_hlg_rtd:*
36: em_hlg_mam:*
37: em_hlg_vrt:*
38: em_sy:*
39: em_hlgo_hum:*
40: em_hlgo_mus:*
41: em_hlgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1737	99.4	1325	9	AF247785 Homo sapi
2	1737	99.4	1380	9	AF348078 Homo sapi
3	1737	99.4	1449	9	BC030948 Homo sapi
4	1737	99.4	1542	6	AX342665 Sequence
5	1725	98.7	1996	6	AR035943 Sequence
6	1713	98.1	90343	9	AC116026 Homo sapi
7	1713	98.1	132745	9	AC068647 Homo sapi
8	1708	97.8	158144	2	AC016455 Homo sapi
9	1231.5	69.6	15543	6	AX376573 Sequence
10	1215.5	69.6	1598	10	AF295367 Mus muscu
11	861	49.3	60298	2	AC116149 Mus muscu
12	727.5	41.6	60298	2	AC116149 Mus muscu
13	504	28.8	3204	10	RN022830 Mus muscu
14	500	28.6	1293	10	MM022829 Mus muscu
15	500	28.6	12630	10	MM022829 Mus muscu
16	500	28.6	288763	2	AC124692 Mus muscu
17	499.5	28.6	2951	5	AF432354 Xenopus l
18	498	28.5	1312	9	HS04950 Homo sapi
19	498	28.5	1426	9	HS042029 Homo sapi
20	498	28.5	1498	5	GDATPREC Homo sapi
21	498	28.5	2424	9	HS042030 Homo sapi
22	498	28.5	16178	9	HS04569 Homo sapi
23	498	28.5	73949	2	AC105795 Rat muscu
24	498	28.5	176170	2	AC021662 Homo sapi
25	498	28.5	176251	9	AC013251 Homo sapi
26	495.5	28.4	1308	4	BT034041 Bos tauru
27	495.5	28.4	1666	4	BT0298CP Bos tauru
28	494	28.3	1180	5	MG009842 Melagris g
29	494	28.3	1192	5	AF012103 Melagris g
30	493.5	28.2	2231	5	AF242850 Raja erin
31	485.5	27.8	3055	9	HSATPRMR Homo sapi
32	485.5	27.8	3055	11	G28604 Homo sapi
33	477	27.3	1014	6	AX148186 Sequence
34	477	27.3	1014	6	AX305130 Sequence
35	477	27.3	1014	6	AX379468 Sequence
36	477	27.3	1014	6	AX384211 Sequence
37	477	27.3	1014	6	AX464561 Sequence
38	477	27.3	1014	6	AB083598 Homo sapi
39	477	27.3	1014	6	AF411109 Homo sapi
40	477	27.3	1081	6	AX458238 Sequence
41	477	27.3	1092	9	AF370886 Homo sapi
42	477	27.3	1414	9	AB065877 Homo sapi
43	477	27.3	9905	6	AX379470 Sequence
44	477	27.3	67645	6	AL356486 Human DNA
45	477	27.3	156555	9	AC026756 Homo sapi

RESULT 1

ALIGNMENTS

AF247785
LOCUS AF247785 1325 bp mRNA linear PRI 26-MAR-2002
DEFINITION Homo sapiens P2Y purinoceptor 1 mRNA, complete cds.
ACCESSION AF247785
VERSION AF247785.1 GI:19716154
KEYWORDS
SOURCE
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS 1 (bases 1 to 1325)
Zhang, W., Li, N., Man, T. and Cao, X.
TITLE Human P2Y purinoceptor 1
JOURNAL Unpublished
2 (bases 1 to 1325)
Zhang, W., Li, N., Man, T. and Cao, X.
REFERENCE
AUTHORS Direct Submission
TITLE Submitted (21-MAR-2000) Department of Immunology, Second Military Medical University & Shanghai Brilliance Biotechnology Institute, 800 Xiangyin Rd., Shanghai 200433, P.R. China
JOURNAL
FEATURES
source
1. .1325
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
69..1073
/product="P2Y purinoceptor 1"
/protein_id="P19716.1"
/db_xref="GI:19716155"
/translation="MLGIMAMNATCKNMALAEALERYKLSIFGIEFVYVGLNTIVVGYSLKMNNSNIYFLNSDIAFLCTPLILSYANGMWYGVDCISRYVLHANTYSLFLFTSIDRLLIKYPEREHLKKKEFLISLAIWVITLPIPLINPYDNGTCCNDPAGSDPNYLLYSKMLTLGFLIPFYMGCFYKIALPLKONROYATAPLEKPLNLYMAVYISVLETPYHMKMNVIASLSWKYQCTOYVINSFYIVRPLAFINSVINPYFYLGDHFMDLMNDLRHFKSLTFSRMAHLLSFRK"
BASE COUNT 359 a 292 c 261 g 413 t
ORIGIN
Alignment Scores:
Pred. No.: 2.94e-132 Length: 1325
Score: 1737.00 Matches: 333
Percent Similarity: 99.70% Conservative: 0
Best Local Similarity: 99.70% Mismatches: 1
Query Match: 99.43% Indels: 0
DB: Gaps: 0
US-09-765-034-2 (1-334) x AF247785 (1-1325)
QY 1 Metleuglyilemetalatrapasnalathrcyslysasntprleualalagluuala 20
DB 69 ATGCTGGGATCATGGCTGGAATGCAACTGGCTGGCAGAGGCTGCC 128
QY 21 Leuglulysrtyrtyrleuserilephetyrlyllegluphevalvalglvalleugly 40
DB 129 CTGAAAAGTACTACCTTTCATTTTATGAGATGAGTTCGTTGGAGGCTCTTGA 188
QY 41 Asnthrilevalvaltyrlyrlyrlyllepheserleuysasntprpasnsersanlle 60
DB 189 AATGCCATGTGTGTTACGGCTACATCTCTCTGAAAGACTGGAACAGCATGTAAT 248
QY 61 Tyrlaupheasnluservalseraspheulaphelueucystlrleupromleuile 80
DB 249 TATCTCTTAACTCTCTCTCTGACTTACCTTTTCTGTCACCCCTCCATCTGATA 308
QY 81 ArgserTyralaasnglyasnrtpiletyrlylvaspyalleucysilesersanrtyr 100
DB 309 AGGAGTATGCAATGGAACATGATATATGAGACGCTCTGCATAGAAGCAACGATAT 368
QY 101 ValleunhisalaasnlueuTyrrhserilleupheleuthrphelieserileasparg 120
DB 369 GTGCTCATGCCAACCTATACAGCATTCCTTCTCATCTTTATACAGCATAGATGA 428

QY 121 Tyrleullellellytyrprophearggluhisleuleuglnlyslgsluphealalle 140
DB 429 TACTGATTAATTAAGTATCTTCCGAGAACCTTTCGAAAAGAGATTGCTATT 488
QY 141 Leulieserleualailetrpvalleualrphleugluleuileuproileuoproleu 160
DB 489 TTAATCTCTTGCCATTGGGTTTATGATACCTTAGAGTTACTACCATCTTCCCTT 548
QY 161 IleasnprovalillethrapsasnlythrhrCysaasnaphealaserSerGlyasp 180
DB 549 ATAAATCTGTATTAACGACAGATGACACCACTTATGATTTTGCAAGTTCTGGAAC 608
QY 181 ProasnrTyraasnlueuiletyrsermetcysleuthrleuleuglypheleuileproleu 200
DB 609 CCCAATCAACCTCATTTACAGCATGTGTCTAACACATGTTGGGGTCTTATTCCTCTT 668
QY 201 PhevalmetCysPhephetrTyrrlysllealeuleupheleuysgluaraspargln 220
DB 669 TTTGTGATGTGTTCTTTTATTAACAGATGCTCTTCTTAAAGCAGAGATAGGAG 728
QY 221 ValalatrhalaleuProleuglulysProleuasnlueualillemetalaivalille 240
DB 729 GTTGCTACTGCTGCTGCCCTTGAAGAGCTCTCAACTTGTCATCATGCGAGTGTATC 788
QY 241 PheserValrProphethrProtyrHisvalmetargasnValargilealaserargleu 260
DB 789 TTCCTGTGCTTTTACACCTCATCATGCGGAGATGAGAGATCGTTCAAGCCTG 848
QY 261 GlyserTPlysglnTyrglncysThrclnValalleasnserserphetrillevalthr 280
DB 849 GGGAGTTGGAAGCATATCATGATGACATGAGTCTGATCAACTCTTTTAAATTTGTGCA 908
QY 281 Argproleualaphelueuasnservalilleasnpovalrphetyrphleuleuglyasp 300
DB 909 CGGCTTTGGCCTTCTGAAACAGTGTCAACCTGCTCTTATTTCTTTGGAGAT 968
QY 301 Hispheargaspmetleuemeasnglueuarghisasnpheylsserleuthrserphe 320
DB 969 CACTTCAGGACATGCTGATTAATCACTGAGACCAACTCAATCCCTTACATCTTT 1028
QY 321 Serargtrpalahisgluleuleuuserpheargglulys 334
DB 1029 AGCAGATGGCTCATGAACTCTTCTTATTCAGAGAAAG 1070
RESULT 2
AF348078
LOCUS AF348078 1380 bp mRNA linear PRI 03-APR-2001
DEFINITION Homo sapiens G-protein coupled receptor 91 (GPR91) mRNA, complete cds.
ACCESSION AF348078
VERSION AF348078.1 GI:13517982
KEYWORDS
SOURCE
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS 1 (bases 1 to 1380)
Wittenberger, T., Schaller, H.C. and Hellebrand, S.
TITLE An expressed sequence tag (EST) data mining strategy succeeding in the discovery of new G-protein coupled receptors
JOURNAL J. Mol. Biol. 307 (3), 799-813 (2001)
MEDLINE 21172992
PUBMED 11273702
REFERENCE
AUTHORS 2 (bases 1 to 1380)
Wittenberger, T., Schaller, C.H. and Hellebrand, S.
TITLE Direct Submission
JOURNAL Submitted (08-FEB-2001) ZMNH, Institut fur Entwicklungsneurobiologie, Martinistr. 52, Hamburg 20246, Germany
FEATURES
source
1. .1380
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="3"

gene /map="3q24-q25.1"
1..1380
/gene="GPR91"
55..1047
CDS /gene="GPR91"
/note="Orphan receptor"
/codon_start=1
/product="G-protein coupled receptor 91"
/protein_id="AAK29080.1"
/db_xref="GI:13517983"

/translation="MAMNATCKMMLAEALAEKYYLIFGYIEFVVGNTIVVGY
/translation="MAMNATCKMMLAEALAEKYYLIFGYIEFVVGNTIVVGY
IFSLKNNMSNIYLFNLSVSDLAFLCTPMLIRVANGNMIVGVLCISNRVLANL
YTSILFEFISIDRLIIRKYPREHLITKKEPALIISAIWVLTLELPLILANL
ITDNGTCNDPAGSGDPKYNLISWCLLGLFLIFPYMCEPYKLAFLKRNQVA
TALPLEKPLNVIMAVVIFSVLETPYHVMNRNRLASRIGSMKQICQVIVNSFIVT
RPLAFNSVINPVEFELGDHFRDMLNMLNRHFKSLTFSFMHELLSFERK"

BASE COUNT 383 a 294 c 274 g 429 t
ORIGIN

Alignment Scores:
Pred. No.: 3.07e-132 Length: 1380
Score: 1737.00 Matches: 333
Percent Similarity: 99.70% Conservative: 0
Best Local Similarity: 99.70% Mismatches: 1
Query Match: 99.43% Indels: 0
DB: Gaps: 0

US-09-765-034-2 (1-334) x AF348078 (1-1380)

OY 1 MetLeuGlyIleMeTAlaTrpPAsnAlaThrCysLysAsnTrpLeuAlaIaGluAlaIa 20
DB 43 ATGCGGGGATCATGCGATGCAATGCACTTGCAGAAATCGCTGCACAGAGCTGCC 102
OY 21 LeuGluLysTyrTyrLeuSerIlePheTyrGlyIleGluPheValaIglValLeuGly 40
DB 103 CTGGAAAGTACTACTCTTCCATTTTATGGATAGACTGGTGGAGACTCTTGA 162
OY 41 AsnThrIleValIValTyrGlyTyrIlePheSerIleuLysAsnTrpAsnSerAsnIle 60
DB 163 AATACCATGTTGGTTTACGGCTACATCTTCTCTGMAAACTGGAAACAGCATATAT 222
OY 61 TyrLeuPheAsnLysSerValSerAspLeuAlaPheLeuCysThrLeuProMetLeuIle 80
DB 223 TATCTCTTAACTCTCTCTCTCTGACTTTCCTGCAACCTCCCGCATGCGATA 282
OY 81 ArgSerTyrAlaAsnGlyAsnTrpIleTyrGlyAspValLeuCysIleSerAsnArgTyr 100
DB 283 AGCAGTATGCGCAATGAAATGCAATATATGAGACGTCTCGCATAGCAACCGATAT 342
OY 101 ValLeuHisAlaAsnLeuTyrThrSerIleLeuPheLeuThrPheIleSerIleAspArg 120
DB 343 GTGCTTCATGCGCAACCTCTATACACACATCTCTCTCTCTCTCTCTCTCTCTCA 402
OY 121 TyrLeuIleIleLysTyrProPheArgLysIleLeuGlnLysLysGluPheAlaIle 140
DB 403 TACTGATATATTAAGATTCCTTCCGAGAACACTTCTGCAAAAGAAAGTTTGTAT 462
OY 141 LeuIleSerLeuAlaIleTrpValLeuValThrLeuGluLeuLeuProIleLeuProLeu 160
DB 463 TTAACCTCTGCGCATTTGGGTTTATGAACTTACTACTACCCATCTCTCTCTCT 522
OY 161 IleAsnProValIleThrAspAsnGlyThrPheCysAsnAspPheIleSerGerglyAsp 180
DB 523 ATTAATCCCTGTTATATACATGACACACCTGTAATGATTTTGCAGAGTCTGAGAC 582
OY 181 ProAsnTyrAsnLeuIleTyrSerMetCysLeuThrLeuLeuGlyPheLeuIleProLeu 200
DB 583 CCCAATCAACCTCATTTAGACAGCATGTGTAAACAGCTGTGGGTTCTTATCTCTCT 642
OY 201 PheValMetCysPhePheTyrTyrIleAlaLeuPheLeuLysGlnArgAsnArgGln 220
DB 643 TTTGTGATGTGTTCTTTATTAACAAGATTGCTCTCTCTCTCTCTCTCTCTCTCTCT 702

OY 221 ValAlaThrAlaLeuProLeuGluLysProLeuAsnLeuValIleMetAlaValAlaIle 240
DB 703 GTTGTACTACTGCTTCCCTTGAAGACCTCTCACTTGTCTATGCGAGTGTATTC 762
OY 241 PheSerValProPheThrProTyrHisValMetArgAsnValArgIleAlaSerArgLeu 260
DB 763 TTTCTGTGTTTATACCCCTATCACTGATCGGAGATGTAGANCGTTCACCCCTG 822
OY 261 GlySerTrpLysGlnTyrGlnCysThrGlnValIleIleAsnSerPheTyrIleValThr 280
DB 823 GGGAGTGGAGACGATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 882
OY 281 ArgProLeuAlaPheLeuAsnSerValIleAsnProValPheTyrPheLeuLeuGlyAsp 300
DB 883 CGGCTTTGGCTTCTTGAACAGTGTATCAACCTGCTCTTATTTCTTTGGAGAT 942
OY 301 HisPheArgAspMetLeuLeuIleAsnGlnLeuArgHisAsnPheLysSerLeuThrSerPhe 320
DB 943 CACTTACGGGACATGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1002

OY 321 SerArgTrpAlaHisGluLeuLeuSerPheArgGluLys 334
DB 1003 AGCAGATGGGCTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1044

RESULT 3
LOCUS BC030948 1449 bp mRNA linear PRI 13-JUN-2002
DEFINITION Homo sapiens, G protein-coupled receptor 91, clone MGC:32514
IMAGE:4594810, mRNA, complete cds.
ACCESSION BC030948
VERSION BC030948.1 GI:21410927
KEYWORDS MGC.

SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 1449)
AUTHORS Strusberg, R.
TITLE Direct Submission
JOURNAL Submitted (03-JUN-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
COMMENT Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: CLONTECH
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mdedpaxll.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
Series: IRAL Plate: 41 Row: e Column: 17
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 14780893.

FEATURES

source
1..1449
/organism="Homo sapiens"
/db_xref="LocusID:56670"
/db_xref="taxon:9606"
/clone="MGC:32514 IMAGE:4594810"
/tissue_type="kidney"
/clone_11b="NIH_MGC_75"
/lab_host="DH10B"
/note="Vector: pDNR-LIB"
100..1104
/codon_start=1

CDS

```
/product="G protein-coupled receptor 91"
/db_xref="GI:21410928"
/translation="MLGIMANATCKNMLAAEALREKYSIFEGIEFVGLNTIV
VGYEISLKNNSNIIYLFNSVDLAFCLPMLIRSYANMNYGVLCISRYLV
HANVTSLIFLFIIDRYLITIKYPERHLLQKEFLALAIWVITLLELPIEL
INPYITDNGTCNDFPASGDPNIIYSMCTILGFLIPLEPKCFYKIALFLKOR
ROVATLPLEKPLNLTMAVIVESVLEFPIHVMNRVIRASRLSKDYQYVINSF
YIVRPLAFINSVINPVFYFLGDHFMDIMNOLHNHKSLSFSEWAEHLISFEREK
BASE COUNT 411 a 308 c 287 g 443 t
ORIGIN
Alignment Scores:
Pred. No.: 3,24e-132 Length: 1449
Score: 1737.00 Matches: 333
Percent Similarity: 99.70% Conservative: 0
Best Local Similarity: 99.70% Mismatches: 1
Query Match: 99.43% Indels: 0
Gaps: 0
US-09-765-034-2 (1-334) x BC030948 (1-1449)
QY 1 Metleuglyllemetalaatrpsanlathrcyslysastrpleuallalaglualala 20
Db 100 ATGCTGGGAGATCATGCAATGCAATGCAAAACTGGCTGCGACAGAGCTGCC 159
QY 21 LeuglulysfyytyleuSerlelpheryrglyilegluphelaValaGlyValleugly 40
Db 160 CTGGAAGAGTACTACCTTCCATTTTATGAGATGAGTCTGCTGGAGATCCTTGCA 219
QY 41 AsanthrilevalaValtyrGlytyrillepheserleuLysAsntrpsanSerSeranile 60
Db 220 AATACCATGTTGTTACGGCTACATCTCTCTGGAAGACTGGAACAGCAGTAATAT 279
QY 61 TytleupheanleuSerValseraspheulaPheluGlythleuprometleuile 80
Db 280 TATCTCTTAACCTCTGCTGCTGACTTGAAGCTTCTGTCGACCCCTCCATGCTGATA 339
QY 81 ArgserfyrAlaasnGlyAsntrpilletyrGlyaspvalleuCyliSerAsnArgrtyr 100
Db 340 AGGAGTATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATG 399
QY 101 ValleuHsilaasnleuTytrTherleleupheleuthrPhelieserlleasparg 120
Db 400 GTGCTTCATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 459
QY 121 TytleuileilelystytrPropheargluHsileuGlnLysleugluphealale 140
Db 460 TACTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 519
QY 141 LeuileSerleuAlailetyrValleuValThleuGluLeuLeuProileuProleu 160
Db 520 TTAATCTCTGCGCATTTGGGTTTAACTTAAGTAACTTAAGTAACTTAAGTAACT 579
QY 161 IleasnProvalIlethAspAsnGlythrcysAsnAspPheliaserSerGlyasp 180
Db 580 ATAATCTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 639
QY 181 ProasntrpsanleuiletyrSermetcysleuthrleuGlyPhleuileProleu 200
Db 640 CCCAGCTCAACCTCATTTACAGCATGCTCAACAGCTGGGGTCTCTTAATCTCTCT 699
QY 201 pheValmetcysphepetyrtyrlyslleAlaleuPhelLysGlnArgAsnArgGln 220
Db 700 TTTTGAGTGTGTTTCTTATTAACAAGATGCTCTCTCCAAACAGAGAAATGGCAG 759
QY 221 ValalatrAlaleuProleuGlyLysProleuAsnleuValIlemetAlaValaIle 240
Db 760 GTTCTACTGCTGCGCCCTGAAAGCCCTCACTTGCTATCATGAGCATGCTGATA 819
QY 241 PheSerValProphetThrProtyrHisValmetArgAsnValArgIleAlaSerArgleu 260
/
product="G protein-coupled receptor 91"
/db_xref="GI:21410928"
/translation="MLGIMANATCKNMLAAEALREKYSIFEGIEFVGLNTIV
VGYEISLKNNSNIIYLFNSVDLAFCLPMLIRSYANMNYGVLCISRYLV
HANVTSLIFLFIIDRYLITIKYPERHLLQKEFLALAIWVITLLELPIEL
INPYITDNGTCNDFPASGDPNIIYSMCTILGFLIPLEPKCFYKIALFLKOR
ROVATLPLEKPLNLTMAVIVESVLEFPIHVMNRVIRASRLSKDYQYVINSF
YIVRPLAFINSVINPVFYFLGDHFMDIMNOLHNHKSLSFSEWAEHLISFEREK
BASE COUNT 411 a 308 c 287 g 443 t
ORIGIN
Alignment Scores:
Pred. No.: 3,46e-132 Length: 1542
Score: 1737.00 Matches: 333
Percent Similarity: 99.70% Conservative: 0
Best Local Similarity: 99.70% Mismatches: 1
Query Match: 99.43% Indels: 0
Gaps: 0
US-09-765-034-2 (1-334) x AX342665 (1-1542)
QY 1 Metleuglyllemetalaatrpsanlathrcyslysastrpleuallalaglualala 20
Db 198 ATGCTGGGAGATCATGCAATGCAATGCAAAACTGGCTGCGACAGAGCTGCC 257
QY 21 LeuglulysfyytyleuSerlelpheryrglyilegluphelaValaGlyValleugly 40
Db 258 CTGGAAGAGTACTACCTTCCATTTTATGAGATGAGTCTGCTGGAGATCCTTGA 317
QY 41 AsanthrilevalaValtyrGlytyrillepheserleuLysAsntrpsanSerSeranile 60
Db 318 AATACCATGTTGTTTACAGCATGCTCAACAGCTGGGGTCTCTTAATCTCTCT 377
QY 61 TytleupheanleuSerValseraspheulaPheluGlythleuprometleuile 80
Db 378 TATCTCTTAACCTCTGCTGCTGACTTGAAGCTTCTGTCGACCCCTCCATGCTGATA 437
QY 81 ArgserfyrAlaasnGlyAsntrpilletyrGlyaspvalleuCyliSerAsnArgrtyr 100
REFERENCE
AUTHORS
Lal, P., Baughn, M.R., Hafalla, A.J., Nguyen, D.B., Gandhi, A.R.,
Kallik, D.A., Griffin, J.A., Yue, H., Khan, F.A., Patterson, C.,
Lu, D.A., Tribouley, C.M., Lu, Y., Wala, N.K., Gaul, R., Yao, M.G.,
Yang, J., Ramkumar, J., Au-Young, J., Hernandez, R., Walsh, R.T. and
Borowsky, M.L.
JOURNAL
Patent: WO 0198351-A 20 27-DEC-2001;
Incyte Genomics, Inc. (US)
FEATURES
source
1..1542
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="Incyte ID No: 3485895CB1"
BASE COUNT 428 a 327 c 315 g 472 t
ORIGIN
Alignment Scores:
Pred. No.: 3,46e-132 Length: 1542
Score: 1737.00 Matches: 333
Percent Similarity: 99.70% Conservative: 0
Best Local Similarity: 99.70% Mismatches: 1
Query Match: 99.43% Indels: 0
Gaps: 0
US-09-765-034-2 (1-334) x AX342665 (1-1542)
QY 1 Metleuglyllemetalaatrpsanlathrcyslysastrpleuallalaglualala 20
Db 198 ATGCTGGGAGATCATGCAATGCAATGCAAAACTGGCTGCGACAGAGCTGCC 257
QY 21 LeuglulysfyytyleuSerlelpheryrglyilegluphelaValaGlyValleugly 40
Db 258 CTGGAAGAGTACTACCTTCCATTTTATGAGATGAGTCTGCTGGAGATCCTTGA 317
QY 41 AsanthrilevalaValtyrGlytyrillepheserleuLysAsntrpsanSerSeranile 60
Db 318 AATACCATGTTGTTTACAGCATGCTCAACAGCTGGGGTCTCTTAATCTCTCT 377
QY 61 TytleupheanleuSerValseraspheulaPheluGlythleuprometleuile 80
Db 378 TATCTCTTAACCTCTGCTGCTGACTTGAAGCTTCTGTCGACCCCTCCATGCTGATA 437
QY 81 ArgserfyrAlaasnGlyAsntrpilletyrGlyaspvalleuCyliSerAsnArgrtyr 100
```

Db 438 AGGAGTTATGCCAATGGAACCTGATATATGAGAGCTGCTCGCTTAAGCAACCATAT 497
 QY 101 VALLLEuHIALAsnuLeuTyThrSerIleuPheLeuThrPheIleSerIleAspArg 120
 Db 498 GTGCTTCATGCACACCTCATATACACGATCTCTTCTTCACTTTATTCAGCATATGATCGA 557
 QY 121 TYrLeuIleIleIleuTyThrProPheArgIleuHisLeuLeuGlnLysGluPheAlaIle 140
 Db 558 TACTGATATATTAAGTATACCTTCCGAGAACACCTTCGCAAAAGAAAGATTGCTATAT 617
 QY 141 LeuIleSerLeuAlaIleTyrrPheValLeuValThrLeuGlnLeuLeuProIleLeuProLeu 160
 Db 618 TTTAACTCTCTTGCCCTTTGGCTTTTAACTTACCTTAACTTACTTACTTCCCTT 677
 QY 161 IleAsnProValIleThrAspAsnGlyThrThrCysAsnAspPheAlaSerSerGlyAsp 180
 Db 678 ATAAATCCGTATATACATGACATGACACACCTGTAATGATTTTGCAGTTCTGAGAC 737
 QY 181 ProAsnTyrrAsnLeuIleTyrrSerMetCysLeuThrLeuLeuGlyPheLeuIleProLeu 200
 Db 738 CCCAATCAACCTCATTTACAGCATGTGTACACACTGTGGGCTTCTTATTCCTCTT 797
 QY 201 PheValMetCysPhePheTyrrTyrrLysIleAlaLeuPheLeuLysGlnArgAsnArgGln 220
 Db 798 TTTTGATGTGTTCTTTTATTAACAGATTGCTCTCTTCTTAAAGCAGAGAAATAGGAG 857
 QY 221 ValAlaThrAlaLeuProLeuGlnLysProLeuAsnLeuValIleMetAlaValAlaIle 240
 Db 858 GTTGCTACTGCTCGCCCTTGAAAAGCCTCACTGATGTCATATGAGCATGAGTGAATC 917
 QY 241 PheSerValProPheThrProTyrrHisValMetArgAsnValArgIleAlaSerArgLeu 260
 Db 918 TTTCTGTGCTTTTACACCTATACGTCATGCGGAATGTAGAGATGCGCTTACACCTG 977
 QY 261 GlySerTrpLysGlnTyrrGlnCysThrGlnValIleIleAsnSerPheTyrrIleValThr 280
 Db 978 GGGAGTGGAAACAGATATAGTCAGCTCAGTCGTCATATACCTCTTTTACATTTGGACA 1037
 QY 281 ArgProLeuAlaPheLeuAsnSerValIleAsnProValPheTyrrPheLeuLeuGlyAsp 300
 Db 1038 CGGCTTTGGCTTCTTCTGAAAGTGTATCAACACCTGCTTCTTATTTCTTTGGAGAT 1097
 QY 301 HisPheArgAspMetLeuMetAsnGlnLeuArgHisAsnPheLysSerLeuThrSerPhe 320
 Db 1098 CACTTCAGGAGCATGTGTATGATCAACAGACACACACTTCAAAACCTTACATCTTT 1157
 QY 321 SerArgTrpAlaHisGlnLeuLeuLeuSerPheArgGlnLys 334
 Db 1158 AGCAGATGGGCTCATGACTCTTACTTTTATTCAGAGAAAG 1199

RESULT 5
 LOCUS AR035943 1996 bp DNA linear PAT 29-SEP-1999
 DEFINITION Sequence 1 from patent US 5871963.
 ACCESSION AR035943
 VERSION AR035943.1 GI:5952611
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 1996)
 AUTHORS Conley,P.B. and Jantzen,H.-M.
 TITLE P2u2 putative receptor and nucleic acid encoding the receptor
 JOURNAL Patent: US 5871963-A 1 16-FEB-1999;
 FEATURES Location/Qualifiers
 source 1..1996
 BASE COUNT 513 a 455 c 381 g 647 t
 ORIGIN

Alignment Scores: 4.3e-131 Length: 1996
 Pred. No.: 1725.00 Matches: 331
 Score:

Percent Similarity: 99.108
 Best Local Similarity: 99.108
 Query Match: 98.748
 DB: 6 Gaps: 0
 US-09-765-034-2 (1-334) x AR035943 (1-1996)
 QY 1 MetLeuGlnIylleMetAlaThrPheAsnAlaThrCysLysAsnTrpPheAlaIleAlaIleAla 20
 Db 625 ATGCTGGGATGCATGGCATGGATGCACTTGGAAAAAAGCTGGGAGAGAGGGCTGCC 664
 QY 21 LeuGlnLysTyrrTyrrLeuSerIlePheTyrrGlyIleGluPheValAlaGlyValLeuGly 40
 Db 685 CTGGAAAGACTACCTTCCATTTTATGGGATGAGTCTGTGGAGAGCTGTGGA 744
 QY 41 AsnTrpIleValValTyrrGlyTyrrIlePheSerLeuLysAsnTrpAsnSerSerAlaIle 60
 Db 745 AATACCATGTTCTTTAGGCTATACATCTCTCTGTGAGAACATGGAACACACTGAATAT 804
 QY 61 TyrrLeuPheAsnLeuSerValSerAspLeuAlaPheLeuCysThrLeuProMetLeuIle 80
 Db 805 TATCTTTAACTCTCTGTCTGTGACTTAACTTTCTGTGGACACCTCCCATGCTGAT 864
 QY 81 ArgSerTyrrAlaAsnGlnAsnTrpIleTyrrGlyAspValLeuCysIleSerAsnArgTyrr 100
 Db 865 AGGAGTTATGCCAATGGAACCTGATATATGAGAGCTGCTGCAATAGCAACCATATAT 924
 QY 101 ValLeuHIALAsnLeuTyrrThrSerIleuPheLeuThrPheIleSerIleAspArg 120
 Db 925 GTGCTTCATGCACACCTCATATACAGCATCTCTTCTTCACTTTATTCAGCATATGCA 984
 QY 121 TYrLeuIleIleIleuTyrrProPheArgIleuHisLeuLeuGlnLysGluPheAlaIle 140
 Db 985 TACTGATATATTAAGTATACCTTCCGAGAACACCTCTGCAAAAGAAAGTTGCTAT 1044
 QY 141 LeuIleSerLeuAlaIleTyrrPheValLeuValThrLeuGlnLeuLeuProIleLeuProLeu 160
 Db 1045 TTAATCTCTTGCCCAATGGGTTTAACTTAACTTAACTTAACTTAACTTAACTTAACT 1104
 QY 161 IleAsnProValIleThrAspAsnGlyThrThrCysAsnAspPheAlaSerSerGlyAsp 180
 Db 1105 ATAAATCCGTATATACATGACATGAGCACACCTGTAAATGATTTTCAAGTTTGGAGAC 1164
 QY 181 ProAsnTyrrAsnLeuIleTyrrSerMetCysLeuThrLeuLeuGlyPheLeuIleProLeu 200
 Db 1165 CCCAATCAACCTCATTTTACAGCATGTGTATCAACCTGTGGGCTTCTTATTCCTCT 1224
 QY 201 PheValMetCysPhePheTyrrTyrrLysIleAlaLeuPheLeuLysGlnArgAsnArgGln 220
 Db 1225 TTTGATGTGTTCTTTTATTAACAGATGTGCTCTTCTTAAAGCAGAGAAATAGGAG 1284
 QY 221 ValAlaThrAlaLeuProLeuGlnLysProLeuAsnLeuValIleMetAlaValAlaIle 240
 Db 1285 GTTGCTACTGCTCTGCCCTTGAAAAGCCTTCACTTGTGTATGAGAGAGTGAATC 1344
 QY 241 PheSerValProPheThrProTyrrHisValMetArgAsnValArgIleAlaSerArgLeu 260
 Db 1345 TTTCTGTGCTTTTAAACACCTATACGTCATGCGGAATGTAGAGATGCTTACGCCCTG 1404
 QY 261 GlySerTrpLysGlnTyrrGlnCysThrGlnValIleIleAsnSerPheTyrrIleValThr 280
 Db 1405 GGGAGTGGAAAGAGATATAGTCAGCTCAGTCGTCATCAACCTCTTTTACATTTGGACA 1464
 QY 281 ArgProLeuAlaPheLeuAsnSerValIleAsnProValPheTyrrPheLeuLeuGlyAsp 300
 Db 1465 CGGCTTTGGCTTCTTCTGAAAGTGTATCAACACCTGCTTCTTATTTCTTTGGAGAT 1524
 QY 301 HisPheArgAspMetLeuMetAsnGlnLeuArgHisAsnPheLysSerLeuThrSerPhe 320
 Db 1525 CACTTCAGGAGCATGTGTATGATCAACAGACACACACTTCAAAACCTTACATCTTT 1584
 QY 321 SerArgTrpAlaHisGlnLeuLeuLeuSerPheArgGlnLys 334

DB 1585 AGCAATGGGCTCATGACTCTACTTTCATTCAGAGAAAG 1626

RESULT 6
AC116026
LOCUS

DEFINITION Homo sapiens 3 BAC RP11-3F11 (Roswell Park Cancer Institute Human BAC library) complete sequence.

ACCESSION AC116026 90343 bp DNA linear PRI 09-APR-2002
VERSION AC116026.1 GI:19697319
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
REFERENCE 1 (bases 1 to 90343)
AUTHORS Mammalia: Eutheria; Primates; Catarrhini; Homnidae; Homo.
Munzy,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C., Alsbrooks,S.L., Amaralunga,H.C., Are,J.R., Ayalew,M., Banks,T., Barbara,J., Benton,J., Blmage,K., Blankenburg,K., Bonnin,D., Buhay,J., Bowie,S., Brieve,M., Brown,E., Brown,M., Bryant,N.P., Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Delaney,K.R., Davis,C., Davy-Carroll,L., Dederich,D.A., Doutharte,K.R., Delgado,O., Deon,A.L., Ding,Y., Dinh,H.H., Earhart,C., Edgar,D., Edwards,C.C., Elha,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Havlik,P., Hawes,A., He,X., Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B., Homsi,F., Howard,S., Huber,J., Huijck,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudan,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtharge,O., Lieu,C., Liu,J., Liu,W., Louissege,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapa,P., Martin,A., Martindale,A., Martinez,E., Massey,E., Mawhinney,E., McLeod,M.P., Meador,M., Mel,G., Metzker,M., Miner,G., Miner,T., Mitchell,T., Mohabadi,K., Moore,S., Morgan,M., Moorish,T., Morris,S., Moser,M., Neal,D., Nelson,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nickerson,E., Nwokenko,S., Oguh,M., Okunolu,G., Oramungu,N., Oyedero,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojuboakan,I., Rolfe,M., Ruiz,S., Saverly,G., Scherer,S., Scott,G., Shen,H., Shooshart,N., Sisson,I., Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Taylor,T., Telitod,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wleczek,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Naylor,S.L., Weinstein,G. and Gibbs,R.

TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 90343)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (23-MAR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE 3 (bases 1 to 90343)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (09-APR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
COMMENT INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

ANNOTATION OF FEATURES:

STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for human and mouse sequences.

Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as low coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: http://gc.bcm.tmc.edu:8088/quality_info/genbank_annotation.html.

QUALSTAT-REPORT.

FEATURES	QUALSTAT-REPORT
source	location/qualifiers
	1. 90343
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/chromosome="3"
	/clone="RP11-3F11"
	991. 1106
repeat_region	/rpt_family="MER45B"
	complement(1314..1627)
repeat_region	/rpt_family="AluSx"
	complement(2137..2430)
repeat_region	/rpt_family="AluY"
	complement(12568..2741)
repeat_region	/rpt_family="L1M4"
	complement(2742..3047)
repeat_region	/rpt_family="AluSx"
	complement(3048..3165)
repeat_region	/rpt_family="L1M4"
	4735..4865
repeat_region	/rpt_family="FLM_C"
	5657..5762
repeat_region	/rpt_family="L1MC/D"
	5906..6237
repeat_region	/rpt_family="LTR21B"
	6289..6773
repeat_region	/rpt_family="HERVFN1"
	complement(8725..9597)
STS	/rpt_family="MER11D"
	12399..12689
repeat_region	/standard_name="136046"
	13774..13816
repeat_region	/rpt_family="Alu"
	13817..13874
repeat_region	/rpt_family="(TA)n"
	complement(15157..15633)
repeat_region	/rpt_family="L2"
	15706..15747
repeat_region	/rpt_family="AT-rich"

repeat_region 16025..16235
/rpl_family="MIR"
repeat_region 16560..16682
/rpl_family="L2"
repeat_region complement(16710..17265)
/rpl_family="LTR49"
repeat_region 18077..18368
/rpl_family="AlusX"
repeat_region complement(18376..18471)
/rpl_family="L2"
repeat_region complement(18486..18859)
/rpl_family="MER57B"
repeat_region complement(20618..20922)
/rpl_family="AlusX"
repeat_region 21337..21363
/rpl_family="AT_rich"
repeat_region 22155..22561
/rpl_family="L1M4"
repeat_region complement(22608..22659)
/rpl_family="L1M4"
repeat_region 22685..23013
/rpl_family="L1M8"
repeat_region 23103..23399
/rpl_family="Alus9"
repeat_region 23500..23973
/rpl_family="L1M3A"
repeat_region complement(24027..24305)
/rpl_family="L1M1"
repeat_region 24304..24655
/rpl_family="L1M3A"
repeat_region 24656..24697
/rpl_family="MADE1"
repeat_region 25203..25518
/rpl_family="AluJo"
repeat_region 25783..25817
/rpl_family="(TAA)n"
repeat_region 26187..26211
/rpl_family="AT_rich"
repeat_region 27014..27030
/rpl_family="AT_rich"
repeat_region complement(27031..27316)
/rpl_family="AlusX"
repeat_region 27317..27328
/rpl_family="AT_rich"
repeat_region 27574..27615
/rpl_family="(TAGA)n"
repeat_region 28062..28166
/standard_name="24707"
STS 28199..28382
/standard_name="13170"
repeat_region complement(29079..29167)
/rpl_family="MT1J"
repeat_region 29168..29532
/rpl_family="THE1B"
repeat_region complement(29533..29552)
/rpl_family="MT1J"
repeat_region 29807..30387
/rpl_family="L1PB2"

Alignment Scores:
Pred. No.: 2,38e-128 Length: 90343
Score: 1713.00 Matches: 328
Percent Similarity: 99.70% Conservative: 0
Best Local Similarity: 99.70% Mismatches: 1
Query Match: 98.05% Indels: 0
DB: 9 Gaps: 0

US-09-765-034-2 (1-334) x AC116026 (1-90343)

OY 6 AlATpASnAlAThTcYsLysAsnTrpLeuAlaIaGlAlaAlaLeuGlnLysTyrTyr 25
|||||
DB 80665 GCATGGAATGCAACTTGGAAAAAATGCTGGCGACGAGGCTGCCCTGGAATAAGTACTAC 80724

OY 26 LeuSerIlePheTyrGlyIleGluPheValIaGlyValLeuGlyAsnThrIleValIa 45
DB 80725 CTTTCATTTTATATGAGATGAGTTCGTTGGAGGCTCCTTGGAATAACATTTGTTGTT 80784
OY 46 TyrGlyTyrIlePheSerIleuAsnTrpAsnSerSerAsnIleTyrIleuPheAsnIleu 65
DB 80785 TACGGCTACATCTCTCTCTGGAAGAACTGGAACAGCAATTAATATCTCTTAACCTC 80844
OY 66 SerValSerAspLeuAlaPheLeuCysThrLeuProMetIleIleArgSerTyrAlaAsn 85
DB 80845 TCTGTCTCGACTTACCTTTCTGTGACACCTCCCATCTGATAGAGATTATGCCAAT 80904
OY 86 GlyAsnTrpIleTyrGlyAspValLeuGlyIleSerAsnArgTyrValLeuHisAlaAsn 105
DB 80905 GGAACCTGATATATGAGAACGTCGTCATACAGCAACCGAATATGCTTCATGCCAAC 80964
OY 106 LeuTyrThrSerIleLeuPheLeuThrPheIleSerIleAspArgTyrIleIleIlys 125
DB 80965 CTTATACACGATCTCTCTCTGCAAGAACTGGAACAGCAATTAATATCTCTTAACCTC 81024
OY 126 TyrProPheArgGlnHisLeuLeuGlnLysGluPheAlaIleLeuIleSerLeuAla 145
DB 81025 TATCCTTCGAGAACACCTTCGCAAAAGAAAGATTGCTATTTTAATCTCTTGCC 81084
OY 146 IleTrpValLeuValThrLeuGluLeuLeuProIleLeuProLeuIleAsnProValIle 165
DB 81085 ATTTGGGTTTATAGTAACTTAGAGTACCAATCTCCCTTAAATCTCTGTTAA 81144
OY 166 ThrAspAsnGlyThrThrCysAsnAspPheAlaSerSerGlyAspProAsnTyrAsnIleu 185
DB 81145 ACTGCAATGAGCACCACTGATATGATTTGCAAGTTGCGAAGCCCACTACCAACCTC 81204
OY 186 IleTyrSerMetCysLeuThrIleuLeuGlyPheIleIleProLeuPheValMetCysPhe 205
DB 81205 ATTTACAGCATGTCCTTAACAGCTGTGGGCTTCCTTATTCCTCTTGTGATGCTTTTC 81264
OY 206 PheTyrTyrIleValIleAlaLeuPheLeuLysGlnArgAsnArgGlnValAlaThrAlaLeu 225
DB 81265 TTTTATATACAGATGCTCTCTCTCCAAAGCAAGAGCAATAGGCAAGTTCCTACTGCTCG 81324
OY 226 ProLeuGluLysProLeuAsnLeuValIleMetAlaValIlePheSerValProPhe 245
DB 81325 CCCCTTGAAAGCCTCTCAACTGTGCATCAATGAGCAGTGTGATATCTCTGCTGCTTTT 81384
OY 246 ThrProTyrHisValMetArgAsnValArgIleAlaSerArgLeuGlySerTrpLysGln 265
DB 81385 ACAACCTATACAGCTCATGCGGAATGAGGATGCTTCACGCTGGGAGTTGGAAGCAG 81444
OY 266 TyrGlnCysThrGlnValValIleAsnSerPheTyrIleValThrArgProLeuAlaPhe 285
DB 81445 TATCAGTGCACGACGCTGCTCATCAACTCTTTTACATTCTGACACGCGCTTGGCCTTT 81504
OY 286 LeuAsnSerValIleAsnProValPheTyrPheLeuLeuGlyAspHisPheArgAspMet 305
DB 81505 CTGAACAGTGCATCAACCTGCTCTCTTCTTCTTTGAGATCACTTACGGGAGCAG 81564
OY 306 LeuMetAsnGlnLeuArgHisAsnPheLysSerIleuThrSerPheSerArgTrpAlaHis 325
DB 81565 CTGATGAATCAACTGAGACACACTTCAATCCCTTACATCCTTTAGCAGATGGGCTCAT 81624
OY 326 GluLeuLeuLeuSerPheArgLys 334
DB 81625 GAACCTCTACTTTCATTCAGAAAAAG 81651

RESULT 7
AC068647
LOCUS: AC068647 132745 bp DNA linear PRI 24-JUL-2002
DEFINITION Homo sapiens 3 BAC RP11-64D22 (Roswell Park Cancer Institute Human
BAC library) complete sequence.
ACCESSION AC068647
VERSION AC068647.10 GI:19774263
KEYWORDS HTG.
SOURCE Homo sapiens.

ORGANISM	TITLE	REFERENCE	AUTHORS
Homo sapiens	Direct Submission	1 (bases 1 to 132745)	Muzny,D.M., Adams,C., Adio-Odunola,B., Ali-osman,F.R., Allen,C., Alsbrooks,S.L., Amarantunge,H.C., Are,J.R., Ayale,M., Banks,T., Barbieri,J., Benton,J., Blinaghe,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowe,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Cartron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhury,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M., Dathorne,S.R., David,R., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinl,H., Douthwalte,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,T., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H.G., Guevara,M., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., He,X., Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B., Homsl,F., Howard,S., Huber,D., Hulyk,S., Hume,T., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jollivet,S., Joudsch,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtenage,O., Lieu,C., Liu,J., Liu,W., Loulegeed,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Maxwell,E., McLeod,M.P., Meador,M., Mel,G., Metker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Moore,S., Morgan,M., Moorishi,T., Morris,S., Moser,M., Neal,D., Nelson,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwunodu,G., Ogunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojoudokun,L., Rolfe,M., Ruiz,S., Savary,G., Scherer,S., Scott,G., Shen,H., Shooshitari,N., Sisson,I., Suttogren,E., Sotnalke,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Sytek,A., Tabot,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,D., Taylor,C., Taylor,T., Tellord,B., Thomas,N., Thomas,S., Usmani,K., Vadequez,L., Vera,V., Villalton,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wlaczek,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Zaylor,S.L., Weinstein,G. and Gibbs,R.
REFERENCE	Direct Submission	2 (bases 1 to 132745)	Worley,K.C.
JOURNAL	Direct Submission	Submitted (06-MAY-2000)	Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE	Direct Submission	3 (bases 1 to 132745)	Worley,K.C.
JOURNAL	Direct Submission	Submitted (26-MAR-2002)	Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE	Direct Submission	4 (bases 1 to 132745)	Worley,K.C.
JOURNAL	Direct Submission	Submitted (28-MAR-2002)	Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE	Direct Submission	5 (bases 1 to 132745)	Worley,K.C.
JOURNAL	Direct Submission	Submitted (29-MAR-2002)	Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE	Direct Submission	6 (bases 1 to 132745)	Worley,K.C.

```

TITLE          Direct Submission
JOURNAL        Submitted (25-JUN-2002) Human Genome Sequencing Center, Department
                of Molecular and Human Genetics, Baylor College of Medicine, One
                Baylor Plaza, Houston, TX 77030, USA
REFERENCE      7 (bases 1 to 132745)
AUTHORS        Morley, K.C.
JOURNAL        Direct Submission
                Submitted (24-JUL-2002) Human Genome Sequencing Center, Department
                of Molecular and Human Genetics, Baylor College of Medicine, One
                Baylor Plaza, Houston, TX 77030, USA
COMMENT        On Mar 20, 2002 this sequence version replaced gi:19718616.
                INFORMATION: http://www.hgsc.bcm.tmc.edu/ or email
                gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the
entire insert of this clone. Overlapping regions of clones are only
sequenced and submitted once, so the sequence for the remainder of
the insert may be found in the record for the adjacent clones.
Overlapping clones are noted at the beginning and end of the
Features listing.

ANNOTATION OF FEATURES:
    STSS are identified using ePCR (Genome Res. 7:541-550) searches
    of a local database that includes entries from dbSTS, GDB, and
    local mapping efforts.
    Repeats are identified using RepeatMasker (A. Smit and P. Green
    unpublished.) for Human and Mouse sequences.
    Genes and Region of sequence similarity are identified by BLAST
    (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the
    EST and cDNA sequences. Genes demonstrate at least two exons
    flanked by consensus splice sites that maintained sequence
    continuity across the splice junctions. Sequences that are not
    identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum
standard of double strand coverage with a minimum of 2 clones and 2
reads with no ambiguities or 2 chemistries with a minimum of 2
clones and 3 reads with no ambiguities. If the sequence quality for
a region does not meet this standard, it will be indicated in the
annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality
standards - estimated error rate less than 1 per 10,000 bases.
Reports of lowest quality individual bases and measures of base
quality are listed below. Description of the metrics can be found
at URL: http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html.

QUALSTAT-REPORT
Location/Qualifiers
1. 132745
  /organism="Homo sapiens"
  /db_xref="taxon:9606"
  /chromosome="3"
  /clone="RP11-64D22"
1. 2005
  /note="Overlaps bases 170209..172213 of clone AC069067"
  /function="clone overlap"
30. 130
  /strand.name="74493"
  complement(522..1015)
  /rpt_family="WL7TD"
  complement(2452..2697)
  /rpt_family="L1MA5A"
  complement(3200..3578)
  /rpt_family="WL7TB"
3600..3749
  /rpt_family="(TA)n"
4391..4411
  /rpt_family="AT_rich"
4909..4960
  /rpt_family="AT_rich"
complement(5657..6403)

```


repeat_region	/rpt_family="L1PA13"	6404..7799	
STS	/rpt_family="L1PA13"	8045..8318	
repeat_region	/standard_name="183647"	complement(8708..9282)	
repeat_region	/rpt_family="L1MD3"	complement(9287..9357)	
repeat_region	/rpt_family="MLT1F1"	complement(9359..9460)	
repeat_region	/rpt_family="L1MC3"	complement(9587..9880)	
repeat_region	/rpt_family="AluSg"	complement(10203..10450)	
repeat_region	/rpt_family="L1MC4"	11468..11699	
repeat_region	/rpt_family="L1M4"	11717..11886	
repeat_region	/rpt_family="MLT2CB"	11908..11982	
repeat_region	/rpt_family="Tigger3(Golem)"	12020..12246	
repeat_region	/rpt_family="THER1C"	12263..12562	
repeat_region	/rpt_family="AluSx"	13326..13346	
repeat_region	/rpt_family="AT_rich"	complement(13934..14245)	
repeat_region	/rpt_family="AluSx"	complement(14256..14568)	
repeat_region	/rpt_family="AluO"	14618..14746	
repeat_region	/rpt_family="MER8"	14825..14849	
repeat_region	/rpt_family="AT_rich"	14865..14906	
repeat_region	/rpt_family="AT_rich"	15465..15739	
repeat_region	/rpt_family="L2"	16579..16756	
repeat_region	/rpt_family="(TTATA)n"	complement(16757..17074)	
repeat_region	/rpt_family="L2"	17621..17660	
repeat_region	/rpt_family="(CAAAA)n"	18544..18725	
repeat_region	/rpt_family="MIR"		
Alignment Scores:			
Pred. No.:	3,59e-128	Length:	132745
Score:	1713.00	Matches:	328
Percent Similarity:	99.70%	Conservative:	0
Best Local Similarity:	99.70%	Mismatches:	1
Query Match:	98.05%	Indels:	0
DB:	9	Gaps:	0
US-09-765-034-2 (1-334) x AC068647 (1-132745)			
QY	6	AlArRpAsnAlArThRcSlySAsnTrPLeuAlAlAgLuAlAlAlAuGlUyLysTyrr	25
Db	123066	GCATGGAGATGCACTTCGACAAATACTGGCTGGCAGACAGAGCTCCCTGGAAAAAGTACTAC	123125
QY	26	LeuSerIlePheTyRgLyIleGluPhVaIValGlYAlLeuGlyAsnThrIleVal	45
Db	123126	CTTTCATTTTTTATGGGATTGAGATTGCTGTGGGACTCCTTGGAATACCATTTGTT	123185
QY	46	TyrGlYrIlePheSerLeuLysAsnTrPAsnSerSeraSnIleTyRleuPheAsnLeu	65
Db	123186	TACGGCTACATCTCTCTGGAAGACGTGAACAGCAGTAATATTATCTCTTAACCTC	123245
QY	66	SeRvAlSerApLeuAlAlPheLeuCySthRleuPrometeuIleatgSerrYrAlAlAsn	85
Db	123246	TCTGTCTGTGACCTAGGCTTTTCTGTGACCCCTCCCATGCTGATTAAGACTTATGCCAAT	123305

QY	86	GIYASNTPRILIEYGLYASPRVALEUCYSILLESERANAGTYVALLENIHISLASN	105
Db	123306	GSAAACSTGGATTTATGGAGACSTGCTGCATACCAACCCGATATGCTGTTATGCCAAC	123365
QY	106	LEUYTTHSRITLIELEUPHEUTHRPHIELSERILIASPARTYUULEULEYLS	125
Db	123366	CTCATACACAGATCTCTCTTCTGACSTTTTACACACATAGACGATCTGATTAATTAAAG	123425
QY	126	TYRPRORHEATRGILHISLEULEGILNYSGLUPHEALAILLEULEISERLEUALA	145
Db	123426	TATCCSTTTCGAGAACACCTTCTGCAAAAAGAAAGAGTTGCTATTTTAAATCCCTGGGCC	123485
QY	146	ILIEPRVALLIEUVALTHLEUNGILUEULEUPROILIELEUPROLEULIASNPROVALILE	165
Db	123486	ATTGTGGTTTATAGTAACTTAAAGTAACTACCCATACCTCCCTTATAAATCCCTGTATA	123545
QY	166	THIASPANGIYTHTRHCYSASNAPSPHEALASERISERGILYASPRDASNTYRASNLEU	185
Db	123546	ACTGACAAATGGGACACACCTGTATATGATTTTGGCAATGTCGGAGACCCCAACACACCTGC	123605
QY	186	ILEYTRSERMETCYSLIEUTHRLEULEUGLYPHELEULIEPROLEUPHEVALMETCYSFHE	205
Db	123606	ATTTCACACATCTGTCATACACTGTGGGGCTTCCTTATCCCTTTTGTGATGTGTTC	123665
QY	206	PHEUTYTYLYLSILIALLEUPHEULEUYSGILNARGASNARGILVALIALATHRALILILEU	225
Db	123666	TTTTATTACAGATTTGCTCTCTCTCTTAAAGCAGAGAAATAGGACAGTTTGCTACTGCTG	123725
QY	226	PROLEUGILYSPROLEUASNLEUVALILEMETALVALIILIEPHESEVALIPROBHE	245
Db	123726	CCCCTTGAAAGCCCTCTCAACTGGTTCATCATGAGGAGTATCTCTCTGCTTTT	123785
QY	246	THRPROYTHHISVALMETARGASNVALARGILIALASERATGLIYSETRPLYSGLN	265
Db	123786	ACACCCTTGATCAGCTCATGCGGAATGTGAGATCGCTTCACGCCCTGGGAGTTGGAAACGAG	123845
QY	266	TYRGINCYSTHGHINVALIILIASNSERPHETRYLILEVALTHTRARGPROLEUNILABHE	285
Db	123846	TATCATGTGCACGACAGTGTCTCATCAACTCTTTACATTTGTATACACGGCTTTTGCTCTTT	123905
QY	286	LEUASNSEVALIILIASNPROVALIPHETRYRPHLEULEUGLYASPHISPEHARGASPHE	305
Db	123906	CTGAACAGATGTCATCAACCCCTGCTCTCATATTTCTTTGGGAGATCAGTCCAGGACATMG	123965
QY	306	LEUMETASNGILNLEUAHGHISANPHELYSSERLEUTHRSEPHSERATGTTRALAHIS	325
Db	123966	CTGAATGATTCAACTAGACACAACTTCAATCCCTTACATCTCTTACGACAGATGGCTCAT	124025
QY	326	GIULEULEULEUSERPHEATRGILYULYS	334
Db	124026	GAACCTCTTACTTTCATTTCAGAGAAAAG	124052
RESULT 8			
LOCUS	AC016455/c	158144 bp	DNA linear HTG 04-MAY-2001
DEFINITION	Homo sapiens clone RP11-3f11, ***	SEQUENCING IN PROGRESS ***	34
ACCESSION	AC016455		
VERSION	AC016455.2	GI:7381788	
KEYWORDS	HTG; HTGS; PHASE1.		
SOURCE	Homo sapiens.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
TITLE	1 (bases 1 to 158144)		
JOURNAL	Birren,B., Linton,L., Nusbaum,C. and Lander,E.		
AUTHORS	Homo sapiens, clone Rp11-3f11		
	Unpublished		
	2 (bases 1 to 158144)		
	Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,		
	Baldwin,J., Barua,N., Beckerly,R., Boguslavsky,L., Boukhalter,B.,		
	Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,		

TITLE
JOURNAL
COMMENT

Cooke, P., DeRellano, K., Dewar, K., Domino, M., Donelan, L., Doyle, M., Ferreira, P., FitzHugh, W., Forrest, C., Funke, R., Gage, D., Galagan, J., Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L., Howland, J., C. Johnson, R., Jones, C., Kann, L., Karacas, A., Klein, J., Lenczky, J., Liu, C., Locke, K., Macdonald, P., Marquis, N., McMan, P., McGurk, A., McKernan, K., McLaughlin, J., Meldrum, J., Morrow, J., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P., Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P., Stange-Thomann, N., Stojanovic, N., Sudirmanian, A., Talamas, J., Tesfaye, S., Titrell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X., Wyman, D., Ye, W.J., Zimmer, A. and Zody, M.

Direct Submission
Submitted (30-NOV-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 1, 2000 this sequence version replaced gi:4479070.
All repeats were identified using RepeatMasker:
Smt, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
Project Information
Center project name: L2722
Center clone name: 3_F_11

NOTE: This is a 'working draft' sequence. It currently consists of 34 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1	1238:	contig of 1238 bp	in length
1339	1338:	gap of 100 bp	
1339	2500:	contig of 1162 bp	in length
2501	2600:	gap of 100 bp	
2601	3814:	contig of 1214 bp	in length
3815	3914:	gap of 100 bp	
3915	5688:	contig of 1774 bp	in length
5689	5788:	gap of 100 bp	
5789	7453:	contig of 1665 bp	in length
7454	7553:	gap of 100 bp	
7554	8459:	contig of 906 bp	in length
8460	8559:	gap of 100 bp	
8560	11388:	contig of 2829 bp	in length
11389	11488:	gap of 100 bp	
11489	14580:	contig of 3092 bp	in length
14581	14680:	gap of 100 bp	
14681	17562:	contig of 2882 bp	in length
17563	17662:	gap of 100 bp	
17663	20871:	contig of 3209 bp	in length
20872	20971:	gap of 100 bp	
20972	24226:	contig of 3255 bp	in length
24227	24326:	gap of 100 bp	
24327	28458:	contig of 4132 bp	in length
28459	28558:	gap of 100 bp	
28559	32034:	contig of 3476 bp	in length
32035	32134:	gap of 100 bp	
32135	34919:	contig of 2785 bp	in length
34920	35019:	gap of 100 bp	
35020	37828:	contig of 2809 bp	in length
37829	37928:	gap of 100 bp	
37929	42243:	contig of 4315 bp	in length
42244	42343:	gap of 100 bp	
42344	44474:	contig of 2131 bp	in length
44475	44574:	gap of 100 bp	
44575	48293:	contig of 3719 bp	in length
48294	48393:	gap of 100 bp	
48394	53132:	contig of 4739 bp	in length
53133	53232:	gap of 100 bp	

	*	53233	57872:	contig of 4640 bp	in length
	*	57873	57972:	gap of 100 bp	
	*	57973	63120:	contig of 5148 bp	in length
	*	63121	63220:	gap of 100 bp	
	*	63221	68799:	contig of 5579 bp	in length
	*	68800	68899:	gap of 100 bp	
	*	68900	74014:	contig of 5115 bp	in length
	*	74015	74114:	gap of 100 bp	
	*	74115	79049:	contig of 4935 bp	in length
	*	79050	79149:	gap of 100 bp	
	*	79150	85094:	contig of 5945 bp	in length
	*	85195	85194:	gap of 100 bp	
	*	85195	90567:	contig of 5373 bp	in length
	*	90568	90667:	gap of 100 bp	
	*	90668	96258:	contig of 5591 bp	in length
	*	96259	96338:	gap of 100 bp	
	*	96339	103095:	contig of 6737 bp	in length
	*	103096	103195:	gap of 100 bp	
	*	103196	110330:	contig of 7135 bp	in length
	*	110331	110430:	gap of 100 bp	
	*	110431	117484:	contig of 7054 bp	in length
	*	117485	117584:	gap of 100 bp	
	*	117585	126459:	contig of 8875 bp	in length
	*	126460	126559:	gap of 100 bp	
	*	126560	134924:	contig of 8865 bp	in length
	*	134929	135024:	gap of 100 bp	
	*	135029	143468:	contig of 8444 bp	in length
	*	143469	143568:	gap of 100 bp	
	*	143569	158144:	contig of 14576 bp	in length
			Location/Qualifiers		
			1. .158144		
			/organism="Homo sapiens"		
			/db_xref="taxon:9606"		
			/clone="RP11-3F11"		
			/clone_lib="RP11-11 Human Male BAC"		
			1. .1238		
misc_feature			/note="assembly_fragment"		
misc_feature			1339. .2500		
misc_feature			/note="assembly_fragment"		
misc_feature			2601. .3814		
misc_feature			/note="assembly_fragment"		
misc_feature			3915. .5688		
misc_feature			/note="assembly_fragment"		
misc_feature			5789. .7453		
misc_feature			/note="assembly_fragment"		
misc_feature			7551. .8459		
misc_feature			/note="assembly_fragment"		
misc_feature			clone_end:SP6		
misc_feature			vector_side:right"		
misc_feature			8560. .11388		
misc_feature			/note="assembly_fragment"		
misc_feature			11489. .14580		
misc_feature			/note="assembly_fragment"		
misc_feature			14681. .17562		
misc_feature			/note="assembly_fragment"		
misc_feature			17663. .20871		
misc_feature			/note="assembly_fragment"		
misc_feature			20972. .24226		
misc_feature			/note="assembly_fragment"		
misc_feature			24327. .28458		
misc_feature			/note="assembly_fragment"		
misc_feature			28559. .32034		
misc_feature			/note="assembly_fragment"		
misc_feature			32135. .34919		
misc_feature			/note="assembly_fragment"		
misc_feature			35020. .37828		
misc_feature			/note="assembly_fragment"		
misc_feature			37929. .42243		
misc_feature			/note="assembly_fragment"		
misc_feature			42344. .44474		
misc_feature			/note="assembly_fragment"		
misc_feature			clone_end:T7		
misc_feature			vector_side:right"		


```

OY 85 AsnGlyAsnTrpIleTyrGlyAspValLeuCysIleSerAsnArgTyrValLeuHisAla 104
DB 314 AATGATTAAGGAGGACCTATGAGATGCTCTCTGTATAAAGCAACCGATATGCTTCACACC 373
OY 105 AsnLeuTyrThrSerIleLeuPheLeuThrPheIleSerIleAspArgTyrLeuIleIle 124
DB 374 AACCTCTACACACGATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 433
OY 125 LysTyrProPheArgGluHisLeuLeuGlnLysGluPheAlaIleLeuIleSerLeu 144
DB 434 AAGTACACCTTCGGAACACCTTCGACAAAAGAGCAATTCGCCATTTAACTCGCG 493
OY 145 AlaIleTyrPheValIleValThrLeuGluLeuLeuProIleLeuProLeuIleAsnProVal 164
DB 494 GCTGCTGGGCTTGTAGTGCTTGTAGTGTACCACTTCCTTCCTTCCTTCCTTCCTTCCTTC 553
OY 165 IleThrAspAsnGlyThrThrCysAsnAspPheAlaSerSerGlyAspProAsnTyrTran 184
DB 554 CCAAAAGAGAGGAGGAGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 613
OY 185 LeuIleTyrSerMetCysLeuThrLeuLeuGlyPheLeuIleProLeuPheValMetCys 204
DB 614 CTCATTTTACAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 673
OY 205 PhePheTyrTyrLysIleAlaLeuPheLeuLysGlnArgAsnArgGlnValAlaThrAla 224
DB 674 TTTCTTCTACTACAGATGCTGCTGCTTAAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 733
OY 225 LeuProLeuGluLysProLeuAsnLeuValIleMetAlaValIlePheSerValPro 244
DB 734 CTGCACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 793
OY 245 PheThrProTyrHisValMetArgAsnValArgIleAlaSerArgLeuGlySerLys 264
DB 794 TTCACACCTTATCATATCATGCGCAATTTGAGAGATCGCCTCAAGCCTGATGTTGGCCA 853
OY 265 GlnTyrGlnCysThrGlnValValIleAsnSerPheTyrIleValThrArgProLeuAla 284
DB 854 CAA---GGATGTACACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 910
OY 285 PheLeuAsnSerValIleAsnProValPheTyrPheLeuLeuGlyAspHisPheArgAsp 304
DB 911 TTTCTGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 970
OY 305 MetLeuMetAsnGlnLeuArgHisAsnPheLysSerIleThrSerPhe 320
DB 971 ATGCTGATTAAGTAAAGTTCAGACAACTCAAGTCCCTTACATTCCTTC 1018

RESULT 11
AC116149 60298 bp DNA linear HTG 25-MAR-2002
LOCUS Mus musculus clone RP24-540E9, LOW-PASS SEQUENCE SAMPLING.
DEFINITION AC116149
ACCESSION AC116149.1 GI:19703273
VERSION HTG; HTGS_PHA5E0.
KEYWORDS Mus musculus.
SOURCE Mus musculus.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 60298)
AUTHORS Birren, B., Linton, L., Nusbaum, C. and Lander, E.
TITLE Mus musculus, clone RP24-540E9
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 60298)
Birren, B., Linton, L., Nusbaum, C., Lander, E., All, A., Allen, N.,
Anderson, S., Batra, N., Bastien, V., Bloom, T., Boguslavsky, L.,
Boukhalter, B., Brown, A., Camarata, J., Campiano, A., Chang, J.,
Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A.,
Cook, A., Cooke, P., Deatellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,
Gibbs, S., Gordon, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagde, B., Horton, L., Hulme, W., Hulme, I., Johnson, R., Jones, C.,
Kamali, A., Karatas, A., Kells, C., Lakocque, K., Lamazares, R.,

```

TITLE JOURNAL COMMENT

Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., MacDonald, P., Major, J., Margulis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Menues, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicoll, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Roselli, M., Roy, A., Santos, R., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Testafaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, M.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Submitted (25-MAR-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIMR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L24912
Center clone name: 540_E_9

* NOTE: This record contains 77 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1 656: contig of 656 bp in length
* 657 756: gap of 100 bp
* 757 1426: contig of 670 bp in length
* 1427 1526: gap of 100 bp
* 1527 2210: contig of 684 bp in length
* 2211 2310: gap of 100 bp
* 2311 2997: contig of 687 bp in length
* 2998 3097: gap of 100 bp
* 3098 3786: contig of 689 bp in length
* 3787 3886: gap of 100 bp
* 3887 4577: contig of 691 bp in length
* 4578 4677: gap of 100 bp
* 4678 5357: contig of 680 bp in length
* 5358 5457: gap of 100 bp
* 5458 6150: contig of 693 bp in length
* 6151 6250: gap of 100 bp
* 6251 6817: contig of 567 bp in length
* 6818 6917: gap of 100 bp
* 6918 7615: contig of 698 bp in length
* 7616 7715: gap of 100 bp
* 7716 8412: contig of 697 bp in length
* 8413 8512: gap of 100 bp
* 8513 9198: contig of 686 bp in length
* 9199 9298: gap of 100 bp
* 9299 9988: contig of 690 bp in length
* 9989 10088: gap of 100 bp
* 10089 10768: contig of 680 bp in length
* 10769 10868: gap of 100 bp
* 10869 11524: contig of 656 bp in length
* 11525 11624: gap of 100 bp
* 11625 12242: contig of 618 bp in length
* 12243 12342: gap of 100 bp
* 12343 13040: contig of 698 bp in length
* 13041 13140: gap of 100 bp

```
* 13141 13829: contig of 689 bp in length
* 13830 13929: gap of 100 bp
* 13930 14647: contig of 718 bp in length
* 14648 14747: gap of 100 bp
* 14748 15451: contig of 704 bp in length
* 15452 15551: gap of 100 bp
* 15552 16247: contig of 696 bp in length
* 16248 16347: gap of 100 bp
* 16348 17028: contig of 681 bp in length
* 17029 17128: gap of 100 bp
* 17129 17802: contig of 674 bp in length
* 17803 17902: gap of 100 bp
* 17903 18593: contig of 691 bp in length
* 18594 18693: gap of 100 bp
* 18694 19375: contig of 682 bp in length
* 19376 19475: gap of 100 bp
* 19476 20082: contig of 607 bp in length
* 20083 20182: gap of 100 bp
* 20183 20875: contig of 693 bp in length
* 20876 20975: gap of 100 bp
* 20976 21650: contig of 675 bp in length
* 21651 21750: gap of 100 bp
* 21751 22427: contig of 677 bp in length
* 22428 22527: gap of 100 bp
* 22528 23238: contig of 711 bp in length
* 23239 23338: gap of 100 bp
* 23339 24028: contig of 690 bp in length
* 24029 24128: gap of 100 bp
* 24129 24803: contig of 675 bp in length
* 24804 24903: gap of 100 bp
* 24904 25603: contig of 700 bp in length
* 25604 25703: gap of 100 bp
* 25704 26357: contig of 654 bp in length
* 26358 26457: gap of 100 bp
* 26458 27140: contig of 683 bp in length
* 27141 27240: gap of 100 bp
* 27241 27946: contig of 706 bp in length
* 27947 28046: gap of 100 bp
* 28047 28734: contig of 688 bp in length
* 28735 28834: gap of 100 bp
* 28835 29536: contig of 702 bp in length
* 29537 29636: gap of 100 bp
* 29637 30324: contig of 688 bp in length
* 30325 30424: gap of 100 bp
* 30425 31130: contig of 706 bp in length
* 31131 31230: gap of 100 bp
* 31231 31910: contig of 680 bp in length
* 31911 32010: gap of 100 bp
* 32011 32691: contig of 681 bp in length
* 32692 32791: gap of 100 bp
* 32792 33482: contig of 691 bp in length
* 33483 33582: gap of 100 bp
* 33583 34274: contig of 692 bp in length
* 34275 34374: gap of 100 bp
* 34375 35081: contig of 707 bp in length
* 35082 35181: gap of 100 bp
* 35182 35861: contig of 680 bp in length
* 35862 35961: gap of 100 bp
* 35962 36660: contig of 699 bp in length
* 36661 36760: gap of 100 bp
* 36761 37447: contig of 687 bp in length
* 37448 37547: gap of 100 bp
* 37548 38243: contig of 696 bp in length
* 38244 38343: gap of 100 bp
* 38344 39034: contig of 691 bp in length
* 39035 39134: gap of 100 bp
* 39135 39813: contig of 679 bp in length
* 39814 39913: gap of 100 bp
* 39914 40597: contig of 684 bp in length
* 40598 40697: gap of 100 bp
* 40698 41392: contig of 695 bp in length
* 41393 41492: gap of 100 bp
* 41493 42190: contig of 698 bp in length
```

```
* 42191 42290: gap of 100 bp
* 42291 42967: contig of 677 bp in length
* 42968 43067: gap of 100 bp
* 43068 43736: contig of 669 bp in length
* 43737 43836: gap of 100 bp
* 43837 44525: contig of 689 bp in length
* 44526 44625: gap of 100 bp
* 44626 45306: contig of 681 bp in length
* 45307 45406: gap of 100 bp
* 45407 46111: contig of 705 bp in length
* 46112 46211: gap of 100 bp
* 46212 46848: contig of 637 bp in length
* 46849 46948: gap of 100 bp
* 46949 47639: contig of 691 bp in length
* 47640 47739: gap of 100 bp
* 47740 48431: contig of 692 bp in length
* 48432 48531: gap of 100 bp
* 48532 49221: contig of 690 bp in length
* 49222 49321: gap of 100 bp
* 49322 50017: contig of 696 bp in length
* 50018 50117: gap of 100 bp
* 50118 50799: contig of 682 bp in length
* 50800 50899: gap of 100 bp
* 50900 51583: contig of 684 bp in length
* 51584 51683: gap of 100 bp
* 51684 52384: contig of 701 bp in length
* 52385 52484: gap of 100 bp
* 52485 53167: contig of 683 bp in length
* 53168 53267: gap of 100 bp
* 53268 53966: contig of 699 bp in length
* 53967 54066: gap of 100 bp
```

Alignment Scores:

```
Pred. No.: 2,31e-59 Length: 60298
Score: 861.00 Matches: 157
Percent Similarity: 85.71% Conservative: 29
Best Local Similarity: 72.35% Mismatches: 31
Query Match: 49.28% Gaps: 0
DB: 2
```

US-09-765-034-2 (1-334) x AC116149 (1-60298)

```
QY 14 TTPleuAlaIaGluaIaAlaLenuCluYsTYrTYrLeuSerIlePheTYrGlyIleGlu 33
DB 3910 TGGGGAATTCAGGCTATCTTGAAATGACCACTCTGCAATTATGACATCGAG 3969
QY 34 PheValValGlyValIleuGlyAsnThrIleValTYrGlyTYrIlePheSerIleuYs 53
DB 3970 TTCAATTTTGGACACGCTTGGGAATGCTCACGTGGTGGCTTCTTCCATGACAG 4029
QY 54 AsnTrpAsnSerAsnIleTYrLeuPheAsnLeuSerValSerAspLeuAlaPheLeu 73
DB 4030 AACTGGAAACACACCAATGCTATCTTTTAACCTTCCACCTTCACTTGGCTTCTC 4089
QY 74 CysThrLeuPrometLeuIleArgSerTYrAlaAsnGlyAsnTrpIleTYrGlyAspAl 93
DB 4090 TGCACCCCTCCCATCTCATTAAGAGTTATGCCATGATGAAGGAGACCTATGAGAGT 4149
QY 94 LeuYsIleSerAsnArgTYrValIleuHisAlaAsnLeuTYrTrsSerIleLeuPheLeu 113
DB 4150 CTCGTATTAACCAACCGATATGCTTCCACCAACCTCTACACGACATCTTCTTCCTC 4209
QY 114 ThrPheIleSerIleAspArgTYrLeuIleIleIleYsTYrPropheArgGluHisLeuLeu 133
DB 4210 ACTTTCAATTAACGATGACCGATATCTGCTCATGAAGTACCTTCCGAGACACTTCTA 4269
QY 134 GlnIysIysGluPheAlaIleuIleSerIleuAlaIleTrpValIleValThrLeuGlu 153
DB 4270 CAAAGAAGGAATTTGCCATTATATCCGCTGCTGTGGGCTTAGTACCTTAGAA 4329
QY 154 LeuLeuProIleLeuProIleAsnProValIleThrAspAsnGlyThrCysAsn 173
DB 4330 GTTCTACCACTGCTCATTCTTATCATTTCTGTCCCAAAAGAGGCGCATTAAGTGCATC 4389
```


Pred. No.:	7 78e-32	Length:	1293
Score:	500.00	Matches:	118
Percent Similarity:	54.49%	Conservative:	64
Best Local Similarity:	35.33%	Mismatches:	114
Query Match:	28.62%	Indels:	38
DB:	10	Gaps:	10
US-09-765-034-2 (1-334) x MMU22829 (1-1293)			
Qy	2	LeuGlyIleMetAlaTrpAsnAlaThrCysLysAsnTrpLeuAlaA	17
Db	92	CTGGCGCGCTTGGGGAACAGTACGTGCTCCCACTGCAGCAGCTTCCCTTCATTC	151
Qy	18	GluaAlaLeuGluys-----TyrTrpLeuSerIlePheTyrGlyIle	32
Db	152	CAATGTGCCCGACCAAGACCGGTTCCAGTTCACCTACCGCGCGCTGCTACATT	211
Qy	33	GluPheValValGlyValLeuGlyAsnThrIleValTyrGlyTyrIlePheSerIle	52
Db	212	GTGTTTCATCATAGCTTCTTCCGACACGCTGCTATCTGATGTTGTTTCCACATG	271
Qy	53	LysAsnTrpAsnSerSerAsnIleTyrLeuPheAsnLeuSerValSerAspLeuAlaPhe	72
Db	272	AAGCTTGGACGGGATCTCCGTGATCATGTTGCTGCTGCTGCTGCTGCTGCTGCTG	331
Qy	73	LeuGlyThrLeuPheMetLeuIleArgSerTyrAlaAsn---GlyAsnTrpIleTyrGly	91
Db	332	GTGTCACCCCTCACAGCCCTCTTCTACATCTTCAACAAAGACTGATCTTCGCG	391
Qy	92	AspValLeuGlyIleSerAsnArgTyrValLeuHisAlaAsnLeuTyrThrSerIleLeu	111
Db	392	GATCTCATGTCACAGCGCAGATTCATCTTCCACGTAATCTCTATGTAAGCATCTTG	451
Qy	112	PheLeuThrPheIleSerIleAspArgTyrLeuIleIleTyrTyrProPheArgLuhis	131
Db	452	TTCCTCACCTGACAGGACACAGGAGTGGCGGTGATACCTCCTCAAGTCTCTG	511
Qy	132	LeuLeuGlnIlysgLuhPheAlaIleLeuIleSerIleAlaIleTyrValLeuValThr	151
Db	512	GCGAGGCTCAAGAAAGAAATGATTTATGTGACGCGTGTGCTGCTGCTGCTGCTG	571
Qy	152	LeuGluLeuLeuProIleLeu-----ProLeuIleAsnProValIle	165
Db	572	GTGGCCATCTCCCTATCTTCTTCTACTCTGCGACTGGGACTCGGAAACAAACTGTC	631
Qy	166	ThrAspAsnGlyThrThrCysAsnAspPheAlaSerGlyAspProAsnTyrAsnLeu	185
Db	632	ACCTGATATGACACCGCTCAATGATATCTGCGAAGT-----TATTTTC	676
Qy	186	IleTyrSerMetCysLeuThrLeuLeuGlyPheLeuIleProLeuPheValMet-----	203
Db	677	ATCTACAGTATGTGACAGCTGTGCGCATGTTCTGCACTCCCTTGGCTGATCTTG	736
Qy	204	CysPhePheTyrTyrIlysgIleAlaLeuPheLeuIlysgIleArgAsnArgGlnValAlaThr	223
Db	737	TGTTATGATTAATTTAAAGCTTTGATTAAATGACCTGCGAAGCTCT-----	787
Qy	224	AlaLeuProLeu---GluysProLeuAsnLeuValIleMetAlaValAlaIlePheSer	242
Db	788	-----CGCGCCGAGGAATATCATTTACTGTTGATTAATTTGCGTACGCTGTTGCT	841
Qy	243	ValProPheThrProTyrHisValMetArgAsnValArgIleAlaSerArgLeuGlySer	262
Db	842	GTGCTTATATCCCTTTCATGATGATGAAACGATGATTTGCGACGCGTTG-----	895
Qy	263	TrpIlysgIlysgIlysgIlysgIlysgIlysgIlysgIlysgIlysgIlysgIlysgIlysg	277
Db	896	-----GATTTCCAGACCCCAAGAAATGTTGATTTCAACGACAGGCTTTATGCCATTTAT	949
Qy	278	IleValThrArgProLeuAlaPheLeuAsnSerValIleAsnProValPheTyrPheLeu	297
Db	950	CAGTACACAGAGGTCTACGACAGTCTGACAGCTGTGTGACCCCATTTCTTATTTCTTG	1009

Qy	298	LeuGlyAspHisPheArgAspMetLeuMetAsnGlnLeuArg	311
Db	1010	GCTGAGATACATTCGAAAGAGACTGTCGAGCCACACAG	1051
RESULT	15		
LOCUS	MMU245636		
DEFINITION	Mus musculus p21 gene for p21 receptor.		
ACCESSION	AJ245636		
VERSION	AJ245636.1 GI:6013074		
KEYWORDS	ADP receptor; P21 gene; P21 receptor.		
SOURCE	Mus musculus.		
ORGANISM	Mus musculus.		
REFERENCE	1 (bases 1 to 12630)		
AUTHORS	Leon, C.		
TITLE	Thymoresistance in p21 receptor knockout mice		
REFERENCE	2 (bases 1 to 12630)		
AUTHORS	Leon, C.		
TITLE	Direct Submission		
JOURNAL	Submitted (04-AUG-1999) LEON C., U311, Inserm, 10 rue Spielmann		
FEATUERS	Submitted (04-AUG-1999) LEON C., U311, Inserm, 10 rue Spielmann		
source	Location/Qualifiers		
	1..12630		
	/organism="Mus musculus"		
	/strain="129/Sv"		
	/sub_species="mus musculus"		
	/db_xref="taxon:10090"		
	/cell_line="ES 129/Sv D3"		
	/dev_stage="embryonic stem cells"		
	/gene="p21"		
	3071..4192		
	/gene="p21"		
	/function="ADP receptor"		
	/codon_start=1		
	/product="p21 receptor"		
	/protein_id="CAB57317.1"		
	/db_xref="GI:6013075"		
	/translation="MTEPWSVYVNGDAEFLAGISLNGNSVYASRAVSSPQCAL		
	TKTGFQYIYPAVYITLITFLGCLNSAIVIMKPYHMKPWGSIYHNLADFLYL		
	TLPALIYFNKDWIFGDAMCKLQRIIFVNLIGSLIFLTCISAHRYSGVVPKLK		
	GRLLKKNIVSVLIVVVAISPILFYSGTGRKRYCYDTSDNDYLSRYFIS		
	MCTVAFCIPILVILGCGYLYALVINDLNSPLRKRKSIYIVIVYFAVSYIPF		
	HVKRTNMLRARLDFOPEMCDENDRVATQVTRGLASLNSCVDPILYFLAGDFRRR		
	LSRATRKASRRSEANLQSKSEMTNLITSEFKONGDRL"		
	4401..4406		
	/gene="p21"		
	4611..4617		
	/gene="p21"		
	9721..9727		
	/gene="p21"		
	10610..10615		
	/gene="p21"		
	10636..10641		
	/gene="p21"		
	11842..11848		
	/gene="p21"		
BASE COUNT	3498 a 2450 c 2763 g 3918 t		1 others
ORIGIN			
Alignment Scores:			
Pred. No.:	8.89e-31	Length:	12630
Score:	500.00	Matches:	118
Percent Similarity:	54.49%	Conservative:	64
Best Local Similarity:	35.33%	Mismatches:	114
Query Match:	28.62%	Indels:	38
DB:	10	Gaps:	10
US-09-765-034-2 (1-334) x MMU245636 (1-12630)			

THIS PAGE BLANK (USPTO)